

#2 D.J. 4-24-02
Priority Papers

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

HYUNG-BAE JEON, ET AL.

For: **SPEECH DETECTION APPARATUS UNDER
NOISE ENVIRONMENT AND METHOD
THEREOF - UTILITY**

11002 U.S. PTO
10/074451
02/11/02

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

Request for Priority

Sir:

Applicant respectfully requests a convention priority for the above-captioned application, namely
Korean application number 01-73007 filed 22 November 2001.

☒ A certified copy of the document is being submitted herewith.

Respectfully submitted,

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN

Dated: _____

2/11/02

Eric S. Hyman, Reg. No. 30,139

12400 Wilshire Blvd., 7th Floor
Los Angeles, California 90025
Telephone: (310) 207-3800

J1002 U.S. PRO

10/074451

02/11/02



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출원번호 : 특허출원 2001년 제 73007 호
Application Number PATENT-2001-0073007

출원년월일 : 2001년 11월 22일
Date of Application NOV 22, 2001

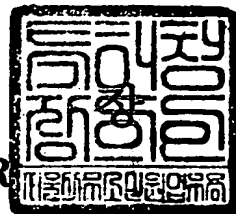
출원인 : 한국전자통신연구원
Applicant(s) KOREA ELECTRONICS & TELECOMMUNICATIONS RESEARCH INST



2001 년 12 월 13 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】 특허출원서
【권리구분】 특허
【수신처】 특허청장
【참조번호】 0004
【제출일자】 2001.11.22
【국제특허분류】 G10L
【발명의 명칭】 잡음 환경에서의 음성신호 검출방법 및 그 장치
【발명의 영문명칭】 Speech detection apparatus under noise environment and method thereof

【출원인】

【명칭】 한국전자통신연구원
【출원인코드】 3-1998-007763-8

【대리인】

【성명】 이영필
【대리인코드】 9-1998-000334-6
【포괄위임등록번호】 2001-038378-6

【대리인】

【성명】 이해영
【대리인코드】 9-1999-000227-4
【포괄위임등록번호】 2001-038396-8

【발명자】

【성명의 국문표기】 전형배
【성명의 영문표기】 JEON, Hyung Bae
【주민등록번호】 760901-1222519
【우편번호】 136-042
【주소】 서울특별시 성북구 삼선동2가 413-17
【국적】 KR

【발명자】

【성명의 국문표기】 정호영
【성명의 영문표기】 JUNG, Ho Young
【주민등록번호】 700308-1696415

【우편번호】	302-150
【주소】	대전광역시 서구 만년동 상아아파트 102동 606호
【국적】	KR
【심사청구】	청구
【취지】	특허법 제42조의 규정에 의하여 위와 같이 출원합니다. 대리인 이영필 (인) 대리인 이해영 (인)
【수수료】	
【기본출원료】	20 면 29,000 원
【가산출원료】	2 면 2,000 원
【우선권주장료】	0 건 0 원
【심사청구료】	12 항 493,000 원
【합계】	524,000 원
【감면사유】	정부출연연구기관
【감면후 수수료】	262,000 원
【첨부서류】	1. 요약서·명세서(도면)_1통

【요약서】**【요약】**

독립 요소 해석 기법의 기저 함수를 이용한 음성 신호 검출 장치 및 그 방법이 개시되어 있다. 본 발명은 소정의 학습 법칙에 의해 음성 신호의 기저함수와 잡음 신호의 기저함수를 훈련하는 과정, 마이크로 입력되는 잡음 특성을 반영하여 상기 과정에서 훈련된 잡음신호의 기저함수를 현재 환경에 적응시키는 과정, 상기 훈련된 음성신호의 기저함수와 상기 환경에 적응된 잡음신호의 기저함수를 바탕으로 음성 구간의 판별정보를 추출하는 과정, 상기 과정에서 추출된 판별정보로 입력되는 마이크 신호의 음성 시작점과 음성 끝점을 검출하는 과정을 포함한다.

【대표도】

도 3

【명세서】**【발명의 명칭】**

잡음 환경에서의 음성 신호 검출 방법 및 그 장치{Speech detection apparatus under noise environment and method thereof}

【도면의 간단한 설명】

도 1은 기저함수들로 선형 조합된 음성 신호의 구성도이다.

도 2는 음성신호를 이용하여 기저함수를 훈련하는 독립요소해석기법 네트워크를 나타내는 개념도이다.

도 3은 본 발명에 따른 음성 신호 검출 장치를 보이는 블록도이다.

도 4는 도 3의 판별정보추출부의 상세도이다.

도 5는 판별정보 추출부에서 추출된 판별정보를 이용하여 음성의 시작과 끝을 판별하는 상태 천이도이다.

도 6은 본 발명에 따른 음성 신호 검출 방법을 보이는 흐름도이다.

【발명의 상세한 설명】**【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<7> 본 발명은 음성 신호 검출 장치 및 그 방법에 관한 것으로서, 특히 독립 요소 해석 기법(Independent Component Analysis)으로 훈련된 기저 함수(basis function)를 이용한 음성 신호 검출 장치 및 그 방법에 관한 것이다.

<8> 일반적으로 음성 인식은 컴퓨터와 같은 장치를 통해서 음성신호를 분석하여 발화된 문장을 출력하는 기술이다. 음성인식 시스템 내에서 음성인식 모듈을 항상 구동하기 위해서는 많은 양의 메모리 등 고비용을 초래하기 때문에 음성이 시작되는 순간에서부터 음성인식 모듈을 동작시켜야 한다. 그렇기 때문에 음성인식 시스템에서 음성구간을 검출 하는 장치는 반드시 필요하게 된다. 또한 음성 구간 검출 방식은 실제 잡음 환경하에서 강인하게 동작해야 하고, 실시간 음성인식기에 쓰이기 위해서 계산량이 적고 실시간으로 동작해야만 한다.

<9> 종래의 음성 구간 검출 장치는 음성 신호의 에너지 성분, 주파수 스펙트럼, 영교차율 정보 등을 이용한다. 그런데 음성 신호에 주변 잡음이 심하게 섞인 경우 음성 신호의 특징들이 잡음에 의해 손상되기 때문에 음성 구간의 검출이 어려워진다. 따라서, 종래의 음성구간 검출 방식은 신호대 잡음비(SNR) 값이 적은, 열악한 잡음 환경에서 음성 검출(Voice Activation Detection) 성능의 저하가 뚜렷하게 나타나며, 그에 따라 묵음을 음성인 것으로 오판하는 허위 경보율(False Alarm Rate)도 높아지게 되는 문제점이 있었다.

【발명이 이루고자 하는 기술적 과제】

<10> 본 발명이 이루고자 하는 기술적 과제는 독립 요소 해석 기법을 이용하여 음성 신호와 잡음 신호의 기저 함수를 학습하고, 학습된 기저 함수를 이용해 신호 대 잡음비가 낮은 고잡음 환경에서도 안정된 음성 구간을 검출하는 음성 신호 검출 방법을 제공하는 데 있다.

<11> 본 발명이 이루고자 하는 기술적 과제는 상기 음성 신호 검출 방법을 적용한 음성 신호 검출 장치를 제공하는 데 있다.

【발명의 구성 및 작용】

- <12> 상기의 기술적 과제를 해결하기 위하여, 본 발명에 의한 잡음 환경 하에서 음성 신호 검출 방법은
- <13> 소정의 학습 법칙에 의해 음성 신호의 기저함수와 잡음 신호의 기저함수를 미리 훈련(training) 하는 과정;
- <14> 마이크로 입력되는 잡음 특성을 반영하여 상기 과정에서 훈련된 잡음신호의 기저함수를 현재 환경에 적응시키는 과정;
- <15> 상기 훈련된 음성신호의 기저함수와 상기 환경에 적응된 잡음신호의 기저함수를 바탕으로 음성 구간의 판별정보를 추출하는 과정;
- <16> 상기 과정에서 추출된 판별정보를 이용하여 음성인식기에 입력되는 마이크 신호의 음성 시작점과 음성 끝점을 검출하는 과정을 포함하는 것을 특징으로 한다.
- <17> 상기의 다른 기술적 과제를 해결하기 위하여, 본 발명에 의한 잡음 환경 하에서 음성 구간을 검출하는 음성 신호 검출 장치에 있어서,
- <18> 소정의 학습 법칙에 의해 음성 신호의 기저함수와 잡음 신호의 기저함수를 훈련하고, 마이크로 입력되는 잡음 특성을 반영하여 훈련된 잡음신호의 기저함수를 현재 환경에 적응시키는 학습네트워크수단;
- <19> 상기 학습네트워크수단에서 훈련된 음성신호의 기저함수와 환경에 적응된 잡음신호의 기저함수를 바탕으로 음성 구간의 판별정보를 추출하는 판별정보추출수단;

- <20> 상기 판별정보추출수단에서 추출된 음성구간의 판별정보로부터 음성인식기에 입력되는 마이크 신호의 음성 시작점과 음성 끝점을 검출하는 음성구간판별수단을 포함하는 것을 특징으로 한다.
- <21> 이하 첨부된 도면을 참조로 하여 본 발명의 바람직한 실시 예를 설명하기로 한다.
- <22> 먼저, 본 발명에서는 잡음에 강한 음성 구간 검출을 구현하기 위해 음성신호와 잡음신호의 기저함수를 이용한다. 기저함수는 음성신호, 또는 잡음신호를 구성하는 요소성분을 말한다. 따라서 기저함수에는 음성신호와 잡음신호의 고유한 특성들, 즉 주파수 특성 등이 반영된다. 기저함수의 이러한 특징을 이용하여, 잡음이 섞인 음성신호에서 잡음과 음성의 상대적인 에너지 비를 구해낼 수 있다.
- <23> 이러한 음성신호와 잡음신호의 기저함수를 구하기 위해서는 독립요소해석기법이 사용된다. 독립요소해석기법은 독립인 신호가 서로 섞여서 마이크를 통해 수집된 신호로부터 원래의 신호가 서로 확률적으로 독립이라는 조건만을 이용하여 혼합되기 전의 신호와 혼합행렬을 찾아내는 방법이다.
- <24> 도 1은 기저함수들로 선형 조합된 음성 신호의 구성도이다.
- <25> 도 1을 참조하면, 음성신호를 x 라고 할 경우, 음성신호(103)는 생성 계수열(s)(101)과 기저함수(102)를 함유하는 혼합행렬(A)에 의해 수학식 1과 같이 구성된다
- <26> 【수학식 1】 $x = As$

- <27> 여기서 혼합행렬(A)의 행벡터가 음성신호의 기저함수가 된다. 독립요소해석기법으로 구한 음성신호의 기저함수(102)는 각각 특정 주파수 성분에 반응하는 파형으로 나타난다.
- <28> 독립요소해석기법의 학습을 통하여 음성신호의 기저함수들을 찾아내는 과정은 다음과 같다.
- <29> 도 2는 음성신호를 이용하여 기저함수를 훈련하는 독립요소해석기법 네트워크를 나타내는 개념도이다.
- <30> 도 2를 참조하면, 독립요소해석기법의 학습 네트워크는 학습데이터로서 충분히 많은 양의 음성 신호를 사용하여 수학적 2에 기초하여 기저함수를 훈련한다.
- <31> **【수학적 2】**
$$\Delta W \propto \frac{\partial H(\hat{u})}{\partial W} W^T W = [I - \Phi(u)u^T]W$$
- <32> 수학적 2와 같은 독립요소해석기법 학습법칙에 의해서 행렬W(202)를 학습하게 되면, 네트워크(W)의 출력 신호인 u(203)는 서로 확률적으로 독립인 신호열이 된다. 신호열(u)은 음성신호(210)중에서 서로 독립인 생성계수 s를 추정된 신호열이다. 반복 학습과정을 수행하여 W가 수렴할 동안 학습을 하여주고, 수렴한 후의 W의 역행렬인 A의 행 벡터가 기저함수가 된다.
- <33> 또한 잡음신호에 대해서도 음성신호와 같이 기저함수들을 학습할 수 있다.
- <34> 그리고 음성 검출을 위해 충분한 양의 음성신호와 다양한 잡음 신호를 이용하여, 미리 음성신호와 잡음 신호의 기저함수를 학습해야한다.
- <35> 도 3은 본 발명에 따른 음성 신호 검출 장치를 보이는 블록도이다.

<36> 도 3을 참조하면, 학습네트웍부(308)는 먼저 충분한 양의 음성신호와 다양한 잡음 신호를 이용하여 독립요소해석기법을 통해 음성신호와 잡음 신호의 기저함수를 미리 훈련(training)하여 메모리등에 저장해둔다. 마이크로부터 신호를 받아 들이는데, 음성 발성이 되기 전 묵음에 해당하는 초기의 음성 인식 대기 상태(301)에서의 마이크 신호에는 현재의 환경에 해당하는 잡음 신호가 들어있다. 음성 인식 대기 상태 동안에 학습네트웍부(308)는 마이크에 입력되는 현재 잡음의 특성을 학습하여 잡음신호의 기저함수(309)를 현재 환경에 맞도록 적응시킨다. 또한 묵음 구간에서의 잡음 특징은 추후 음성 시작점과 음성 끝점을 판별하기 위해 사용할 문턱값을 조정하는데 사용한다.

<37> 음성구간판별부(310)는 판별정보추출부(303)에서 추출된 판별정보를 바탕으로 음성시작점과 음성끝점을 판별한다. 음성구간판별부(310)를 상세하게 설명하면, 마이크 신호(302)가 입력되면 판별정보 추출부(303)는 이미 학습된 음성신호의 기저함수와 학습네트웍부(308)에서 현재 환경에 맞도록 적응된 잡음신호의 기저함수를 이용하여 음성 시작 및 끝점을 판별하는 판별정보를 계산한다. 음성 시작점판별부(304)는 판별정보추출부(303)에서 추출된 판별정보를 사용하여 음성 시작점을 검출한다. 음성 인식부(305)는 음성시작점판별부(304)로부터 음성이 시작되었다는 정보가 넘어오면 마이크 신호(302)로부터 음성인식을 수행한다. 음성 끝점 검출부(306)는 판별정보추출부(303)에서 받은 판별정보와 음성인식부(305)의 인식결과를 이용하여 마이크 신호 중 음성 신호가 끝나는 지점을 검출한다. 이때 음성 시작점 판별부(304)와 음성 끝점 판별부(306)는 상태 천이 알고리즘에 의해 음성구간을 판별한다.

- <38> 음성의 끝점을 검출 한 후 다시 음성 인식 대기 상태(307)로 복귀하며, 이때 학습네트워크부(308)에서 현재 환경의 잡음 특성과 판별 문턱값의 적응이 이루어 진다.
- <39> 도 4는 도 3의 판별정보추출부(303)의 상세도이다.
- <40> 도 4를 참조하면, 먼저 학습네트워크부(308)는 독립요소해석기법을 통하여 미리 음성신호의 기저함수(408)와 잡음신호의 기저함수(409)를 갖는다. 음성 신호(301)가 들어오면 음성기저함수계수추출부(402)는 음성 기저함수(408)를 이용하여 음성 생성 계수를 추정한다. 이 추정된 음성생성계수는 각 기저함수가 음성신호에 얼마나 기여하고 있는지를 의미하는 양을 나타낸다. 또한 잡음기저함수계수추출부(403)는 잡음신호 기저함수(409)를 이용하여 역시 잡음신호의 생성계수를 추정한다.
- <41> 음성 가능성도 계산부(404)는 추정된 음성 생성 계수를 파라미터로 하여 음성신호일 가능성을 의미하는 음성신호 가능성도(likelihood)를 계산한다.
- <42> 잡음 가능성도 계산부(405)는 추정된 잡음 생성 계수를 파라미터로 하여 잡음 신호일 가능성을 의미하는 잡음신호 가능성도(likelihood)를 계산한다. 본 발명에서는 가능성도(likelihood)에 로그(log)를 취한 로그가능성도(log-likelihood)를 사용한다.
- <43> 음성신호의 로그가능성도(log-likelihood)는 수학식 3과 같이 계산된다.
- <44> **【수학식 3】** $\log p(x | \theta) = \log p(s) - \log(\det | A_s |)$

<45> 여기서 x 는 마이크 신호이며, θ 는 파라미터(기저함수, 생성계수 등)이며, s 는 음성신호(speech)이며, A_s 는 음성기저함수 정보를 갖는 혼합행렬이다.

<46> 또한 잡음 신호의 로그가능성도(log-likelihood)는 수학식 4와 같이 계산된다.

<47> **【수학식 4】** $\log p(x | \theta) = \log p(n) - \log(\det | A_n |)$

<48> 여기서 x 는 마이크 신호이며, θ 는 파라미터(기저함수, 생성계수 등)이며, n 은 잡음신호(noise)이며, A_n 는 잡음기저함수 정보를 갖는 혼합행렬이다.

<49> 판별정보계산부(406)는 음성 가능성도 계산부(404) 및 잡음 가능성도 계산부(405)에서 계산된 가능성도(likelihood)값을 이용하여 음성 시작점과 끝점 판별에 이용할 파라미터 정보(406)들을 계산한다. 기본적으로 묵음 구간에서는 음성신호와 잡음신호의 로그가능성도(log-likelihood)값들이 비슷하였다가, 음성구간에서는 음성 신호의 로그가능성도(log-likelihood)의 값이 크게 증가하므로, 음성 및 잡음신호의 로그가능성도(log-likelihood)값들의 차이를 판별 정보로 이용한다.

<50> 음성 시작점을 찾는 판별정보I은 다음과 같이 구한다. 즉, 음성신호의 로그가능성도(log-likelihood)와 잡음신호의 로그가능성도(log-likelihood)와의 차이를 초기 묵음구간에서의 음성 로그가능성도와 잡음 로그가능성도의 차이로 정규화하여 판별정보로 사용한다. 그리고 이와 함께 잡음신호의 로그가능성도(log-likelihood)값이 음성의 고주파 성분에 잘 반응하는 특징을 이용하여 잡음

신호의 로그가능성도(log-likelihood)를 정규화된 두 로그가능성도 (log-likelihood)의 차와 함께 사용하여 최종적으로 판별정보I을 구한다.

<51> 음성의 끝점을 찾는 판별정보II는 다음과 같이 구한다. 즉, 일정시간 동안의 음성구간에서의 음성신호와 잡음신호의 로그가능성도(log-likelihood)간 차의 변동폭을 음성시작점의 두 로그가능성도 (log-likelihood)의 차이값으로 정규화하여 판별정보로 사용한다. 이 판별정보는 음성이 끝나고 나서 다시 묵음이 시작하게 되면 두 로그가능성도 (log-likelihood)의 변동폭이 적은 값으로 수렴하는 성질을 사용한 것이다. 정규화된 두 로그가능성도의 차와 함께 음성인식기의 인식결과를 이용하여 최종적으로 판별정보 II를 구하여 준다.

<52> 도 5는 판별정보 추출부(303)에서 추출된 판별정보 I과 판별정보 II를 이용하여 음성의 시작과 끝을 판별하는 상태 천이도이다.

<53> 먼저, 초기의 잡음이 혼합된 묵음상태(501)에서 마이크 신호가 입력된다. 판별정보 I이 문턱값 I을 넘으면 시작점 대기상태(502)로 이동한다. 이어서, 잡음 환경에 둔감한 성능을 내기 위하여 일정 시간 이상 시작점 대기 상태(502)에 머물러야 음성 구간 상태(503)로 천이하도록 하였다. 이때 일정시간(Num I)을 카운트하기 위해 Count I을 사용한다. Count I은 초기상태(501)에서 0으로 초기화된다. 시작점 대기상태(501)에서 판별정보 I이 문턱값 I보다 크게 되면 시작점 대기상태(501)에 머무르게 되는데, 이 때 Count I을 1씩 증가시켜주어 일정 시간동안 시작점 대기상태(502)에 머물렀는지를 체크하게 된다. Count I이 정해진 기준값 Num I 보다 크게 되면, 즉 일정 시간 이상 시작점 대기 상태(502)에 머물렀다면 음성구간 상태(503)로 이동한다. 시작점 대기 상태에서 음성구간상태로 천이

한 순간에서 Num I 이전의 시점이 음성 신호의 시작점이 된다. 시작점 대기 상태(502)에 머무르는 동안 판별정보 I이 문턱값 I보다 작아서 다시 초기 상태(501)로 이동 한다면, 시작점 대기 상태(502)에 머무르는 시간을 기록하고 있는 Count I은 0으로 다시 초기화된다.

<54> 음성구간상태(503)에서 판별정보 II가 문턱값 II보다 크게 되면 음성 구간 상태(503)에 머무른다. 음성구간상태(503)에서 판별정보 II가 문턱값 II보다 작게 되면 끝점대기상태(504)로 이동한다. 이어서, 끝점대기상태(504)에서 판별정보 II가 문턱값 II보다 작게 되어야 끝점대기상태(504)에 머무르게 되고, 머무르는 시간이 일정 시간 이상 되어야만 초기묵음상태(501)로 이동하게 된다. 머무르는 시간은 Count II로 세어준다. 끝점 대기 상태에서 초기 묵음 상태로 천이한 순간에서 Num II 이전의 시점이 음성 신호의 끝점이 된다. 끝점대기상태(504)에 머무르는 동안 판별정보 II가 문턱값 II보다 크게 되면 음성구간상태(503)로 되 돌아간다. 음성구간상태(503)로 이동하면 Count II는 0으로 초기화 된다.

<55> 이어서, 음성의 끝점을 검출하여 끝점 대기 상태(504)에서 초기묵음상태(501)로 이동하게 되면 다시 음성의 시작점 검출을 수행한다. 이때 초기묵음상태(501)에서 판별정보 I이 문턱값 I보다 작은 동안 계속 머무른다.

<56> 도 6은 본 발명에 따른 음성 신호 검출 방법을 보이는 흐름도이다.

<57> 먼저, 음성을 포함하는 마이크 신호가 입력된다(602과정).

<58> 이어서, 마이크 신호로부터 생성계수를 추정하고(603과정), 추정한 생성계수와 기저함수로부터 가능성도(likelihood)값을 계산한다(604과정).

- <59> 이어서, 음성신호의 가능성도(likelihood)값과 잡음신호의 가능성도(likelihood)값으로부터 판별정보I를 계산한다(605과정).
- <60> 이어서, 판별정보I로부터 음성 시작점이라고 판정되면(606 과정) 마이크 신호는 음성 신호 구간이 된다.
- <61> 이어서, 음성이 시작되면 마이크 신호를 음성인식기에 입력하고(608 과정), 다른 한편으로 음성 끝점 검출을 위해 음성 신호로부터 생성계수를 추정하고(609 과정), 추정한 생성계수와 기저함수로부터 가능성도(likelihood)값을 계산한다(610과정).
- <62> 이어서, 구해진 음성신호의 가능성도(likelihood)값과 잡음신호의 가능성도(likelihood)값으로부터 음성 끝점 판정을 위한 판별정보II를 계산한다(611과정).
- <63> 이어서, 판별정보II에 의해 음성 끝점이라고 판정되면(613 과정) 음성신호로부터 시작점과 끝점을 찾아낸다.(613과정).
- <64> 반면에, 음성신호가 입력되지 않는 초기의 음성인식 대기 상태인 잡음이 섞인 묵음 구간에서는 학습을 통해 잡음신호 기저함수를 현재 잡음 환경에 맞도록 적응시키며, 또한 현재 잡음의 상태에 따라 시작점과 끝점을 결정할 때 사용하는 문턱값I,II를 적응시킨다(607 과정).
- <65> 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포

함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 하드디스크, 플로피디스크, 플래쉬 메모리, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드로서 저장되고 실행될 수 있다.

【발명의 효과】

<66> 상술한 바와 같이 본 발명에 따른 잡음 환경에서의 음성 신호 검출 방법 및 그 장치에 의하면 독립 요소 해석 기법으로 훈련된 기저 함수(basis function)를 이용함으로써 잡음이 심한 환경에서도 음성신호를 오류 없이 찾아낼 수 있다. 또한 계산량도 기존의 방법에 비해 적기 때문에 실시간 시스템에의 활용이 가능하다. 따라서 실시간으로 동작하는 음성인식기의 경우 잡음 환경하에서도 강인하게 음성신호를 찾아내어 인식기의 성능을 개선할 수 있다.

【특허청구범위】**【청구항 1】**

잡음 환경 하에서 음성 신호 검출 방법에 있어서,

소정의 학습 법칙에 의해 음성 신호의 기저함수와 잡음 신호의 기저함수를
훈련(training) 하는 과정;

마이크로 입력되는 잡음 특성을 반영하여 상기 과정에서 훈련된 잡음신호의
기저함수를 현재 환경에 적응시키는 과정;

상기 훈련된 음성신호의 기저함수와 상기 환경에 적응된 잡음신호의 기저
함수를 바탕으로 음성 구간의 판별정보를 추출하는 과정;

상기 과정에서 추출된 판별정보로 입력되는 마이크 신호의 음성 시작점과
음성 끝점을 검출하는 과정을 포함하는 음성 신호 검출 방법.

【청구항 2】

제1항에 있어서, 상기 소정의 학습 법칙은 독립요소해석기법임을 특징으로
하는 음성 신호 검출 방법.

【청구항 3】

제1항에 있어서, 상기 판별 정보 추출 과정은

상기 음성 기저함수와 잡음 기저함수를 바탕으로 음성 및 잡음 생성 계수
를 추정하는 과정;

상기 과정에서 추정된 음성 및 잡음 생성 계수로 음성 신호 가능성도와 잡
음 신호 가능 성도를 계산하는 과정;

상기 과정에서 계산된 음성 신호 가능성도값과 잡음 신호 가능성도값 간의 차이로 음성 구간 판별 정보를 계산하는 과정을 포함하는 것을 특징으로 하는 음성 신호 검출 방법.

【청구항 4】

제3항에 있어서, 상기 음성 신호 가능성도는

$$\log p(x | \theta) = \log p(s) - \log(\det | A_s |) \quad \text{로 계산되며,}$$

여기서 x 는 마이크 신호이며, θ 는 파라미터이며, s 는 음성 신호이며, A_s 는 음성기저함수 정보를 갖는 혼합 행렬임을 특징으로 하는 음성 신호 검출 방법.

【청구항 5】

제1항에 있어서, 상기 과정에서 음성 시작점을 찾는 판별정보는 음성 신호의 가능성도와 잡음 신호의 가능성도간의 차이를 초기 묵음구간에서 음성신호의 가능성도와 잡음신호의 가능성도간의 차이로 정규화한 값을 특징으로 하는 음성 신호 검출 방법.

【청구항 6】

제1항에 있어서, 상기 과정에서 음성 시작점을 찾는 판별정보는 음성 신호의 가능성도와 잡음 신호의 가능성도간의 차이를 초기 묵음구간에서 음성신호의 가능성도와 잡음신호의 가능성도간의 차이로 정규화한 값과 잡음 신호의 가능성도를 사용하는 것을 특징으로 하는 음성 신호 검출 방법.

【청구항 7】

제1항에 있어서, 상기 과정에서 음성 끝점을 찾는 판별정보는 일정시간 동안의 음성구간에서의 음성 및 잡음 가능성도의 차의 변동폭을 음성시작점의 음성 및 잡음 가능성도의 차이값으로 정규화한 값임을 특징으로 하는 음성 신호 검출 방법.

【청구항 8】

제1항에 있어서, 상기 음성 시작점과 음성 끝점 검출 과정은

잡음이 혼합된 초기묵음상태에서 마이크 신호가 입력되면 음성시작점 판별 정보가 제1문턱값을 넘으면 시작점 대기상태로 이동하며, 일정 프레임동안 음성 시작점판별정보가 제1문턱값 보다 크면 음성구간 상태로 이동하고 그렇지 않을 경우 초기 묵음 상태로 복귀하며, 상기 음성구간에서 음성 끝점 판별정보가 제2 문턱값보다 적으면 음성 끝점 대기상태로 이동하며, 상기 끝점 대기상태에서 일정 시간 이상 존재하면 초기 묵음 상태로 이동하고 그렇지 않을 경우에는 음성구간으로 복귀하는 것을 특징으로 하는 음성 신호 검출 방법.

【청구항 9】

제8항에 있어서, 상기 제1 및 제2 문턱값은 현재 잡음 상태에 따라 결정되는 것임을 특징으로 하는 음성 신호 검출 방법.

【청구항 10】

잡음 환경 하에서 음성 구간을 검출하는 음성 신호 검출 장치에 있어서,

소정의 학습 법칙에 의해 음성 신호의 기저함수와 잡음 신호의 기저함수를 혼련하고, 마이크로 입력되는 잡음 특성을 반영하여 상기 잡음신호의 기저함수를 현재 환경에 적응시키는 학습네트워크수단;

상기 학습 네트워크수단에서 혼련된 음성신호의 기저함수와 환경에 적응된 잡음신호의 기저함수를 바탕으로 음성 구간의 판별정보를 추출하는 판별정보추출수단;

상기 판별정보추출수단에서 추출된 음성구간의 판별정보로 입력되는 마이크 신호의 음성 시작점과 음성 끝점을 검출하는 음성구간판별수단을 포함하는 음성 신호 검출 장치.

【청구항 11】

제10항에 있어서, 상기 판별정보추출수단은

상기 음성 기저함수를 이용하여 음성 생성 계수를 추정하는 음성기저함수 계수추출부;

상기 잡음신호 기저함수를 이용하여 잡음신호의 생성계수를 추정하는 잡음 기저함수계수추출부;

상기 음성기저함수계추출부에서 추정된 음성 생성 계수로 음성신호가능성도를 계산하는 음성 가능성도 계산부;

상기 잡음기저함수계추출부에서 추정된 잡음 생성 계수로 하여 잡음신호가능성도(likelihood)를 계산하는 잡음 가능성도 계산부;

상기 음성가능성도계산부 및 잡음가능성도계산부에서 계산된 음성 신호 가능성도값과 잡음 신호 가능성도값간의 차이로 음성 판별 정보를 계산하는 판별정보계산부를 포함하는 것을 특징으로 하는 음성 신호 검출 장치.

【청구항 12】

프로세서를 구비한 컴퓨터 시스템에 있어서,

소정의 학습 법칙에 의해 음성 신호의 기저함수와 잡음 신호의 기저함수를 미리 훈련하는 과정;

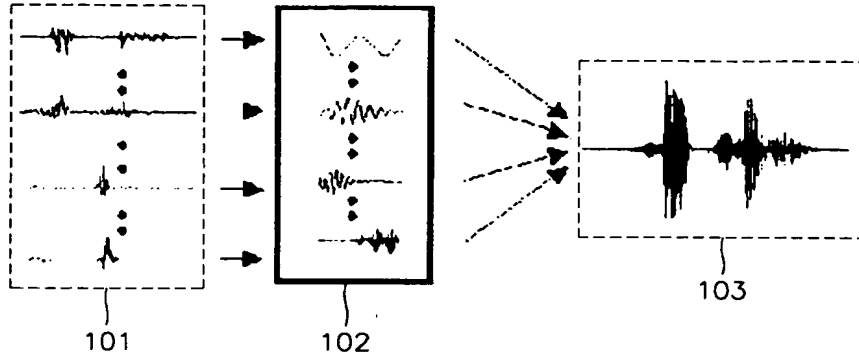
마이크로 입력되는 잡음 특성을 반영하여 상기 과정에서 훈련된 잡음신호의 기저함수를 현재 환경에 적응시키는 과정;

상기 훈련된 음성신호의 기저함수와 상기 환경에 적응된 잡음신호의 기저함수를 바탕으로 음성 구간의 판별정보를 추출하는 과정;

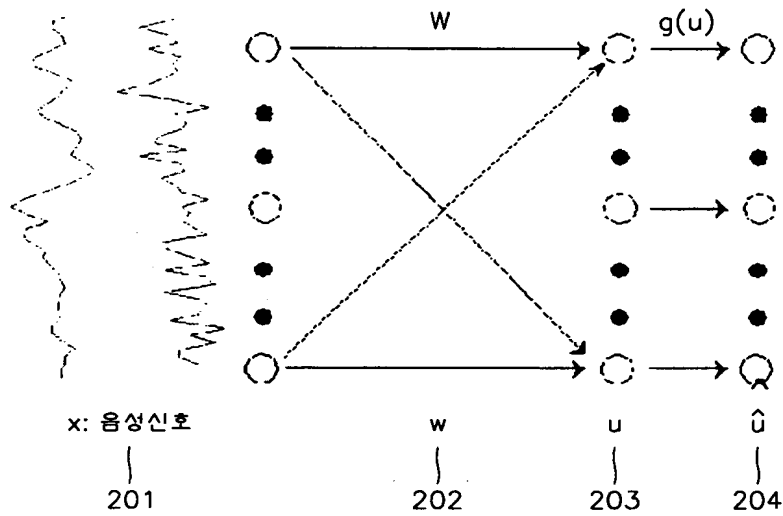
상기 과정에서 추출된 판별정보로 입력되는 마이크 신호의 음성 시작점과 음성 끝점을 검출하는 과정을 포함하는 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록 매체.

【도면】

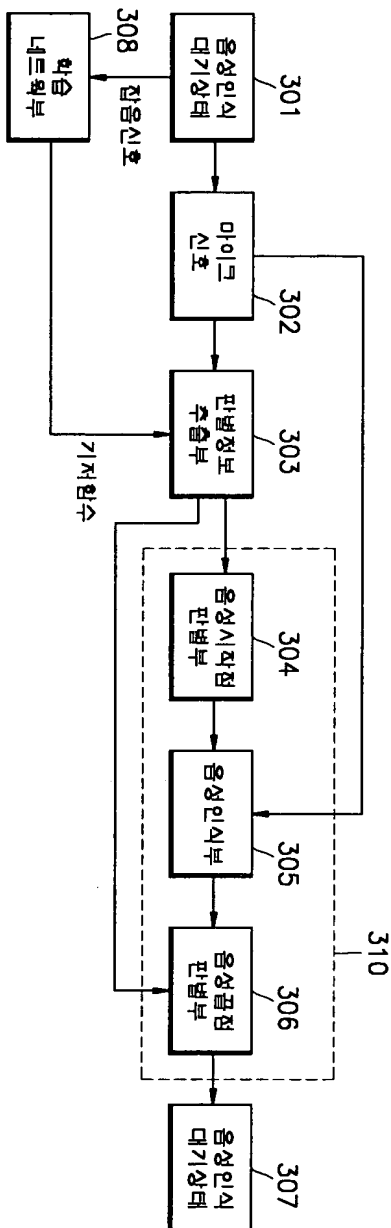
【도 1】



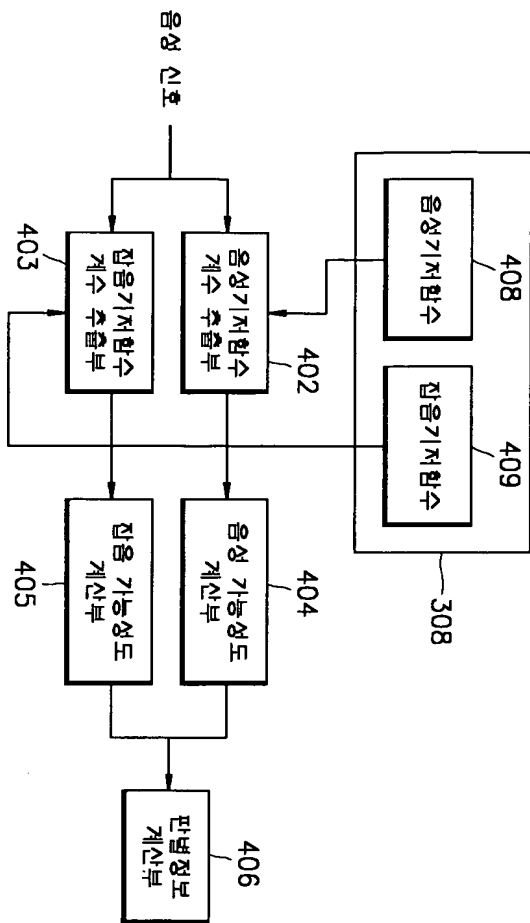
【도 2】



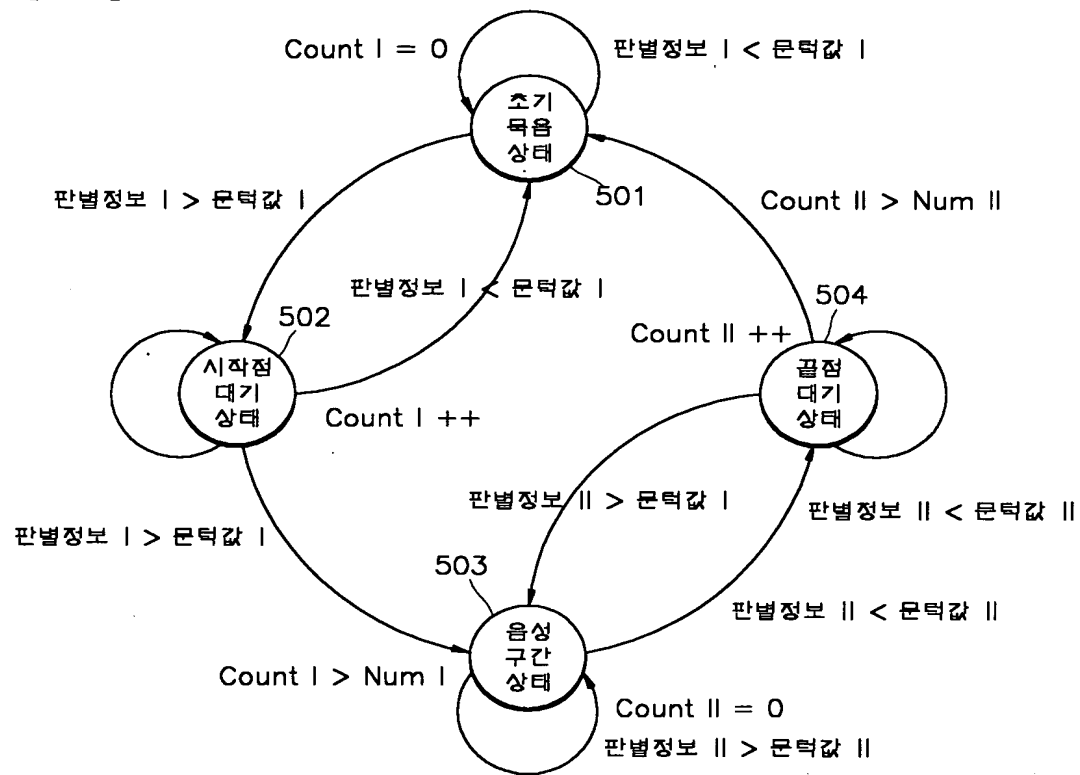
【도 3】



【도 4】



【도 5】



【도 6】

